

A Roadmap for Natural Language Processing Research in Information Systems

Dapeng Liu
Virginia Commonwealth
University
Liud22@vcu.edu

Yan Li
Claremont Graduate University
yan.li@cgu.edu

Manoj A. Thomas
Virginia Commonwealth
University
mthomas@vcu.edu

Abstract

Natural Language Processing (NLP) is now widely integrated into web and mobile applications, enabling natural interactions between human and computers. Although many NLP studies have been published, none have comprehensively reviewed or synthesized tasks most commonly addressed in NLP research. We conduct a thorough review of IS literature to assess the current state of NLP research, and identify 12 prototypical tasks that are widely researched. Our analysis of 238 articles in Information Systems (IS) journals between 2004 and 2015 shows an increasing trend in NLP research, especially since 2011. Based on our analysis, we propose a roadmap for NLP research, and detail how it may be useful to guide future NLP research in IS. In addition, we employ Association Rules (AR) mining for data analysis to investigate co-occurrence of prototypical tasks and discuss insights from the findings.

1. Introduction

Natural Language Processing (NLP) is an interdisciplinary field of computer science, artificial intelligence, and linguistics that explores how computers can be used to understand and manipulate natural language text or speech. Although NLP is considered as a maturing academic research area with its own cumulative history since 1950s [1], it has mostly remained as an area of research within the computational linguistic domain. Traditionally, most NLP related research were conducted at universities and funded research institutions. However, the vastness of textual data through electronic communication systems, social media, and World Wide Web combined with the need for quick access to specific and comprehensive information has driven the advancement and commercial adoption of NLP in

recent years. Nowadays, NLP is widely integrated with web and mobile applications, enabling natural interactions between human and computers. It has matured to the point where spoken language is portrayed as the next human-computer interface [2].

As a discipline, Information Systems (IS) research investigates the complementary sides of technology and social science. Although NLP is often a sub-system integrated with an application, its growing presence surely cannot be ignored. In IS, NLP research generally concerns the refinement and application of NLP techniques to solve real-world problems [3], such as creating spoken dialogue systems [4], speech-to-speech translation engines [5], mining social media [6] for individual information analysis, or identifying sentiment toward products and services [7]. However, researchers in IS have yet to establish NLP as a mainstream research area. A comprehensive review of IS NLP scholarship is thus desired to understand what role IS plays in the overall NLP landscapes and how NLP research fits within the overall IS research agenda.

In this paper, we first conduct a thorough review of IS literature to assess the current state of NLP research. The initial assessment indicates that most research has thus far focused on building and evaluating design artifacts to solve NLP-related problems. We synthesize 12 prototypical NLP tasks based on the in-depth analysis of NLP literature, and propose a roadmap for IS NLP research. Guided by the roadmap, we analyze 238 NLP articles published in IS journals between 2004 and 2015. We then apply association rules (AR) mining to gain additional insights on the interplay between different prototypical tasks commonly used in NLP applications. The contribution of this research is threefold. First, based on an in-depth review of literature, we identify and categorize 12 prototypical NLP tasks that are widely researched. Although published research individually contributes to exploring one or more NLP tasks and related techniques, none have comprehensively reviewed or

synthesized tasks most commonly addressed in NLP research. Our synthesis would thus benefit researchers and practitioners interested in NLP. Second, we propose a roadmap for NLP research in IS by emphasizing its design science orientation. We describe how NLP research may be conducted using the roadmap. Third, based on our analysis, we highlight that, although NLP is recognized as a significant area of IS research, the lack of behavioral research in this important domain remains to be addressed.

The rest of the paper is organized as follows. We first provide a background on NLP (section 2), followed by the initial assessment of NLP research in IS (section 3). We propose a roadmap to guide NLP research in IS (section 4), following which we analyze data and summarize findings (section 5). We conclude by highlighting key considerations for future research (section 6).

2. Background

Natural language (NL) refers to any human written or spoken language that has evolved naturally for human communication. The interaction between computers and human generally comprises two branches of activities: NL understanding and NL generation. NL understanding concerns the computational process of transferring natural language collected from human to a machine understandable format. NL generation focuses on computer systems that can produce understandable texts in human language. While NL understanding and NL generations share similar theoretical foundations and are used together in many real world applications, the internal process of these two activities are quite different [8]. Essentially, NL understanding is the process of mapping human language into computational representation (i.e., given some NL inputs, how to choose an appropriate interpretation among multiple possible ones), and NL generation is the process of mapping computational representation into human language (i.e., given the different means to achieve the desired NL outputs, how to decide which one to use) [8].

Both NL understanding and NL generation are concerned with computational models of language, which requires essential linguistic understanding of all aspects of language [9], including words and parsing, parts of speech (POS) and morphology (word formation), phrases (word order) and grammars, lexical and sentence semantics, syntactic or semantic ambiguities, phrase structure, etc. NLP research has evolved from empirical-based

approaches (i.e. based on the physical symbol system hypothesis) to statistic-based, where quantitative approaches, such as machine learning algorithms, are adopted to facilitate automated language processing. Examples of machine learning algorithms that have been applied in NLP research include genetic algorithm (GA) [10], Naive Bayes [11], decision tree (DT), support vector machine (SVM) [12], hidden Markov model (HMM) [13], entropy model [14], etc.

Once NL is processed into machine-readable formats, various tasks can be carried, such as information extraction [15, 16] and spoken language processing [17]. While it is not practical to list an exhaustive set of NLP-related tasks, based on a thorough review of literature, we synthesize 12 prototypical NLP tasks: text classification or categorization (TC), information extraction (IE) and information retrieval (IR), semantic annotation, text summarization (TS), machine translation (MT), corpus analysis, text generation, sentimental analysis or opinion mining, NL inference (NLI), grammatical text analysis, word sense disambiguation (WSD), and speech recognition.

As an applied discipline, NLP research in IS often seeks to solve real world problems by applying existing NLP-related algorithms and tasks [3]. The research output may include several design artifacts implementing one or more prototypical task and/or algorithms. For example, Valencia-Garcia et al. [18] designed a system to translate surgeon's natural language into robot-executive commands by integrating speech recognition, information extraction, semantic annotation, and inference tasks.

3. Initial Assessment of NLP Research

To understand the current state of NLP research in IS, we conducted a three step assessment of NLP literature. First, to create a complete NLP literature corpus, we followed Webster and Watson [19] to search scholarly databases using terms such as "Natural Language Processing," and "NLP" in title, abstract, or keywords in IS journals from 2004 to 2015. Databases searched included EBSCOHost's Academic Search Complete and Business Source Complete, all databases within Proquest (e.g., ABI/INFORM Complete), and all databases within PsycNet (e.g., PsycINFO). We then reviewed every retrieved article to make sure they were NLP related publications. Our search identified 238 journal articles for further analysis. The initial assessment of retrieved literature shows an increasing trend in NLP research, especially since 2011 (Figure 1).

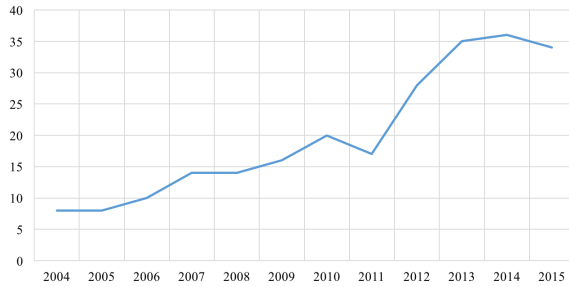


Figure 1. Count of articles by year

Second, to determine how NLP research fits within IS, we conducted a detailed assessment of methodologies commonly used in NLP research by comparing 12 methodologies in use and applicable to MIS research compiled by Palvia et al. [20]. Our review identified some article employing one or more methodologies. In such cases, we coded each article up to four methodologies. The initial coding result shows that the top four methodologies that most commonly appear in NLP research are framework or model or method or instantiation (88.66%), laboratory experiment (84.87%), secondary data (21.43%) and library research (7.14%). The remaining eight methodologies account for less than 5% of published NLP literature, four of which (i.e., interview, field study, qualitative research, and field experiment) did not appear in any NLP literature.

Surprisingly, Palvia et al. [20] did not include design science research in their categorization, whereas our NLP literature corpus is dominated by design science articles. Design science research commonly includes the representation of design artifacts (e.g., framework, model, or instantiation) and the evaluation of artifacts (e.g., experiment, questionnaire). In our in-depth review, we therefore coded each design science article based on its design artifact and its evaluation method. However, in many instances, this approach posed a problem during our coding process. For example, Leopold et al. [21] designed an NLP-based method to automatically detect naming convention violations, and instantiated the method into a java-based prototype system. To demonstrate the utilities of the designed method, an experimental environment was set up where document collections were analyzed using both the prototype system and manual inspection benchmark. While this research is clearly a design science research, it does not directly fit any categories proposed by Palvia et al. [20]. We therefore coded the study into two categories, “frameworks or models or method or instantiation” per its design artifact, and “laboratory experiment” per its evaluation. Ideally, this study should have been categorized as design

science with subcategories of artifact types and evaluation methods. Given the growing prominence of design science research in IS, its inclusion as an important research methodology is critical. Thus, an update to Palvia et al. [20] categorization to include design science is highly warranted.

Table 1: Summary of keyword frequency in NLP literature

Rank	Keyword	Freq.	%
1	Ontology	19	7.95%
2	Information retrieval	16	6.69%
3	Text mining	15	6.28%
4	Information extraction	11	4.60%
5	Machine learning	10	4.18%
6	Sentiment analysis	9	3.77%
7	Algorithms	9	3.77%
8	Opinion mining	7	2.93%
9	Knowledge acquisition	7	2.93%
10	Question answering	6	2.51%
11	Artificial intelligence	6	2.51%
12	Experimentation	5	2.09%
13	Named entity recognition	5	2.09%
14	Data mining	5	2.09%
15	Word sense disambiguation	5	2.09%

Third, we analyzed keyword frequencies in our document corpus, assuming that keywords in a journal article would highlight important NLP topics studied in the given research. NLP related tasks are an integral part of NLP applications that process spoken language and text. Identifying common NLP tasks in IS research can provide us a good understanding of the NLP research landscape. Table 1 ranks frequencies of each keyword and its percentage in total keywords. The keyword that appears most frequently is “ontology”, which is a formal specification of a shared conceptualization [22], and the building block for inference techniques based on semantic web technologies. A closer look at related literature shows other ontology-related tasks such as NL inference [23] or semantic annotation [24]. This highlights popularity in research that integrates advanced and powerful NLP techniques with ontology-based machine readable domain vocabulary. Other frequently appearing keywords such as “information retrieval,” “information extraction,” “sentiment analysis”, “opinion mining”, “word sense disambiguation”, are all representative of prototypical NLP task we synthesized. Additionally, three closely related keywords, “text

mining”, “machine learning”, and “data mining”, together represent one of the most important application areas of NLP, i.e., the process of extracting interesting and non-trivial patterns from unstructured text documents. Almost all text mining research in our literature corpus involved the use of NLP methods and techniques. NLP research in IS journals also cover conventional application areas such as “knowledge acquisition” and “artificial intelligence”. The increasing discussion of “algorithms” and “experimentation” indicates that NLP research is gaining maturity and credence among IS researchers.

4. Roadmap for NLP research in IS

Based on above discussion, it is evident that existing frameworks for reviewing IS literature do not fit the NLP research scheme, which is dominated by design science. Hevner et al. [25] proposed a conceptual framework for understanding, executing, and evaluating design science research in IS. Additionally, Hevner [26] presented three closely related cycles of activities as an embodiment of design science research. Relevance cycle is the gathering of requirements from the environment;

design cycle is the building and evaluating of design artifacts and design process; and rigor cycle is the grounding of design efforts in the knowledge base and contributing to broaden the knowledge base. In this research, we align with Hevner et al. [25] and Hevner [26], and propose a roadmap for NLP related design science research in IS as shown in Figure 2. The road map includes three levels: lexical sources, IS research and knowledge base. The dotted lines connect two or more prototypical NLP tasks with real world artifacts. Below, we describe each level and illustrate how three levels interact with each other when conducting NLP research in IS.

Lexical sources are an integration of all aspects of language [9], including words and parsing, POS and morphology, sentences and phrases, grammar, and semantics, etc. It provides the environment for NLP and represents the problem space of NLP research. Lexical sources function in the relevance cycle delimiting the scope of NLP research in IS. As the foundation of NLP, lexical sources serve at the lowest level. Basic linguistic processing (tokenization, POS tagging, co-reference resolution, syntactic parsing, etc.) is often required to prepare the lexical sources for analysis.

Knowledge base is the foundational body of

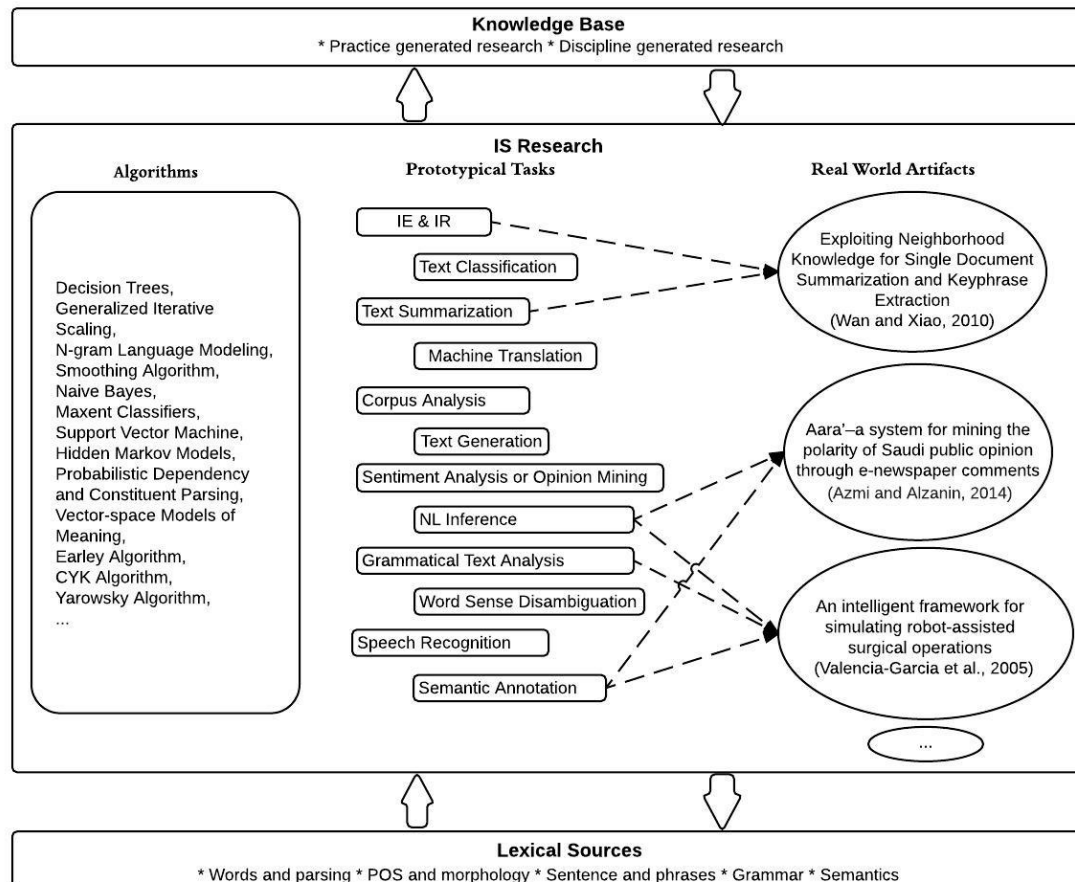


Figure 2: Roadmap for NLP research in IS

knowledge and methodologies that provides the raw material from and through which IS research in design science may be accomplished [25]. In Figure 2, the knowledge base is represented at the top illustrating its function of providing “foundational theories, frameworks, instruments, constructs, models, methods, and instantiations used in the develop/build phase” [25] for NLP research. The knowledge base for NLP research serves in the rigor circle, and assists the selection and application of appropriate methods, techniques, and theories in the construction and evaluation of NLP artifacts. Based on our extensive literature review, we categorize NLP research in IS into two areas namely, practice generated and discipline generated. Practice generated research investigates NLP design artifact (algorithms, methods, NLP-based systems, etc.) and their impact on human agency. For example, Demirtas et al. [27] contribute to the practice generated research by investigating and proposing two methods of automatic categorization and summarization of documentaries using subtitles of videos. Discipline based research is the understanding and abstracting from phenomena of interest and the artifact. For example, Arazy and Woo [28] provided the preliminary evidence for the usefulness of NLP techniques for IR by investigating the effect of three key parameters (i.e., directionality, distance, and weighting) on collocation indexing performance.

NLP-related design science research in IS, in the middle level of Figure 2, focus on three dimensions, namely algorithms, prototypical NLP tasks, and real world artifacts. Both lexical sources and IS knowledge base serve as research inputs. Research outputs contribute back to the knowledge base and/or lexical sources. Based on our literature review, we propose three inter-related dimensions to NLP-related design science research in IS.

Dimension 1: Algorithms

Algorithms have received considerable attention from scholars in NLP research. NLP related algorithms can be divided into two groups: rule-based and statistics-based. The former emphasizes rules that are abstracted and inducted from the characteristics of language, while the latter focuses on the use of statistical methods to analyze large scale of corpora. The adoption of statistical NLP methods and machine learning algorithms may lead to purposeful design artifacts that benefit from NL understanding, generation, and utility. For example, machine learning algorithms, such as GA, SVM, and HMM are now widely employed to implement many prototypical NLP tasks (e.g., computer assisted text generation using IE and IR techniques [29],

classification of sentiment [30]). Advancements in algorithms and computational methods may also facilitate better performance of specific NLP tasks (e.g., speech recognition [31]).

Dimension 2: Prototypical tasks

In NLP research, a design artifact targeting a real world problem may combine several prototypical tasks. Bird et al. [32] provided an explicit description of three prototypical tasks (i.e. text classification, IE, and grammar-based analysis). Mills and Bourbakis [33] surveyed graph-based techniques for NLP tasks such as classification, semantic similarity analysis, and IE. Similarly, Montoyo et al. [34] surveyed sentiment analysis, and Karimi et al. [35] explored the state of art of machine translation. From the perspective of machine learning, Wong et al. [36] analyzed ontology based reasoning and inference. Although these studies individually contribute to exploring one or more NLP tasks and related techniques, none have comprehensively reviewed or synthesized the tasks most commonly addressed in NLP research. Through our analysis of the literature, we identify and categorize 12 prototypical NLP tasks (section 2) that are widely researched.

Dimension 3: Real world artifacts

Design research in IS targets the development of purposeful artifacts targeting unsolved real world problems [37, 25]. Correspondingly, the evaluation of artifact for its functionality, completeness, consistency, accuracy, performance, reliability, usability and fit are integral to design science research. Since NLP research leans strongly towards design science (section 2) and address important real world issues, it is imperative that NLP artifacts are evaluated appropriately. The implementation of a NLP artifact designed to solve a real world problem often combines multiple prototypical tasks. For instance, to address the challenge of classifying twitter messages into topics, Vilares et al. [38] presents a framework that includes a linguistic knowledge extractor (IE) and a topic classifier (text classification). For IS researchers, this demonstrates how NLP algorithms may be applied to NLP tasks in the development of real world applications.

5. Data Analysis and Findings

Guided by the proposed Roadmap, we coded all articles into three dimensions (i.e., algorithms, prototypical NLP tasks, and real world artifacts). We then analyzed these articles based on coding results and synthesized observations that NLP researchers and practitioners may benefit from.

214 of the 238 articles cover one or more prototypical tasks described in section 2, most of

which utilize one or more basic NLP algorithms. For example, Arazy and Woo [28] attempted to enhance information retrieval by applying standard vector space model to calculate text collocation indexes. Yazdani and Popescu-Belis [39] proposed a method for computing semantic relatedness and applied the method to tasks of semantic annotation, IR, and text classification. Their method improves upon graph-based random walk algorithms for NLP problems. Similarly, a new approach for multi-document summarization [40] utilizes hierarchical Bayesian models. Some studies used multiple prototypical tasks to solve real world problems. For example, Wierzbicki et al. [41] used sentiment analysis and TC to improve the computational trust representation in existing trust management systems. Valencia-Garcia et al. [18] presented an intelligent framework for simulating robot-assisted surgical operations that employed semantic annotation and inference for simulating surgical operation.

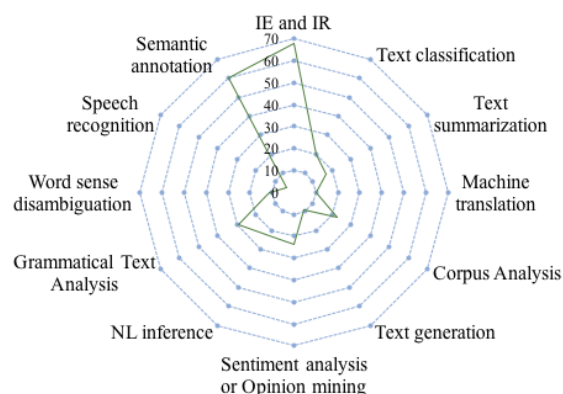


Figure 3. Prototypical tasks commonly studied in NLP research

Among the 24 articles that did not fit one of the NLP tasks, two articles focused on the improvement of NLP-related algorithms: first is a semi-supervised learning algorithm that combined co-training with active learning [42], and the second is an improved conditional random field model for NLP sequential data modeling that incorporated temporal dependencies between variables. Two articles surveyed existing NLP-related algorithms (Mills and Bourbakis [33] surveyed graph-based methods, Al-Shawakfa et al. [43] compared Arabic root finding algorithms). The rest 20 articles focused on real world NLP-related applications, though no prototypical tasks were explicitly represented. For example, Rus et al. [44] reported advances in intelligent tutoring systems with conversational

dialogue, where each system could employ many complex NLP related tasks but not explicitly stated.

The radar chart in Figure 3 illustrates the frequency (actual number of articles) with which the 12 prototypical tasks appear in our literature corpus. It can be seen that tasks such as “IE and IR” and “semantic annotation” have received the most attention, which we discuss in detail below.

The popularity of IE and IR research lies in the exponential growths of unstructured data on the Web and the ubiquity of NLP and machine learning. IE is a process to extract specific information (including entities, relationships, events, etc.) from natural language text [45]. The abstracted data then can be processed and stored into structured form for future analysis. IE usually requires heavy NLP and semantic inference. IR is a deeper field that involves NLP, statistics, artificial intelligence (AI), and data science, where mathematical models are applied to the text corpora to discover patterns in the data and retrieve the information needed [25]. While IR evolved from searching different form of computerized content, a principle driver of modern IR innovations has been the explosion of information published from tens of millions of content creators on the World Wide Web. It compels to better annotate and analyze billions of web pages so users can quickly search for information that is both relevant and comprehensive to their needs [46]. Generally, IE and IR are not clearly separable. For example, concept level extraction, a subtask of IE, applies IR models along with AI-based inference rules to extract related concepts from text. The ability to develop IE dynamically, in part, relies on advancements in IR modeling. The popularity of AI engines, such as IBM Watson that utilizes IE and inferencing, has casted a spotlight on both IE and IR fields. In IS, IE and IR attract strong interests from researchers not only to develop IR and IE models and engines [47, 48], but also to extract and store knowledge from NL for other prototypical tasks [49, 50, 51]. The literature also shows broad real world applications for IE and IR, such as a tool for mining Wikipedia [52], automatic case acquisition from texts [53], creating virtual human dialogue [4], computer assisted writing systems [29], etc.

Research in semantic annotation has received much attention as the growth of textual web calls for the semantic web, where formal structure and semantics are added to the web content for more efficient information management and access. In a nutshell, semantic annotation is the task of linking class and instance information about entities in the text to their semantic descriptions [24]. Semantic annotation requires a basic ontology (or taxonomy) to

define entity classes, entity identifiers to distinguish entities and link their semantic descriptions, and a knowledge base to store these entity descriptions. The popularity of semantic annotation in NLP research may reflect upon the highest frequency of “ontology” as a keyword (Table 1). Most research related to semantic annotation adopts ontology to represent semantic knowledge in NL [49, 54]. Semantic annotation can also be employed to achieve interoperability and eliminate heterogeneity in texts [55]. Our analysis highlights the synergy between NLP research and web semantics. In fact, fundamental tasks of building semantic web include both semantic annotation (formally annotate and hyperlink entities in textual web documents) and IR (index and retrieve documents with respect to annotation) [24]. This also reflects upon our proposed roadmap, where broader applications of NLP research to solve real world problems involve two or more prototypical tasks.

In our coding, we found that majority of articles included more than one prototypical tasks. To investigate if the presence of one task would imply the presence of other tasks in the same article, we employed Association Rules (AR) mining for data analysis. AR mining is a popular pattern discovery method in knowledge discovery and data mining (KDDM). It was first introduced by Agrawal et al. [56] to mine large transactional databases. The objective of AR mining is to find items that imply the presence of other items (prototypical tasks in our case) in the same transaction (i.e. each article can be viewed as a transaction). It can be expressed as $A \Rightarrow B$ (e.g., bread \Rightarrow Peanut Butter & Jelly), where A and B are sets of items in a given transaction t_i , and $A \Rightarrow B$ meets both the minimal *support* and minimal *confidence* constraints. *Support* specifies the probability that a transaction t_i contains both items, A and B. *Confidence* specifies the conditional support, given that the transaction already contains A. It should be noted that an AR does not always imply causation. Both *support* and *confidence* constraints are probability-based measures. In our data analysis, we first set the minimal confidence level as 25%, resulting in 5 AR rules. In addition, we want to find associated prototypical tasks that are departing from independence and positively correlated, for which we use *lift*, a measure of departure from independence [57]. A *lift* value greater than 1 implies that A and B appear more frequently together than expected under independence, and vice versa. Thus, only four rules are considered in our analysis (Figure 4), for which we next provide detailed discussions by reflecting back on our literature corpus.

Lift	Confidence (%)	Transaction Count	Rule
2.90	27.27	3.00	WSD \Rightarrow Corpus Analysis
1.39	44.44	4.00	Text Generation \Rightarrow IE_IR
1.22	35.00	7.00	Corpus Analysis \Rightarrow Semantic Annotation
1.11	31.82	7.00	NL Inference \Rightarrow Semantic Annotation

Figure 4: AR mining result

WSD involves the association of a given word in a text or discourse with a definition (sense), which is distinguishable from other meanings potentially attributable to that word [34]. Most WSD studies we reviewed focused on the co-occurrences of words to measure context similarity. Corpus such as WordNet, serves as a useful frame of reference when implementing WSD. The strong association between the WSD and corpus analysis in our AR analysis indicates that developing algorithms and methods to minimize corpus WSD (i.e., precise concept identification) continues to be a popular NLP research area. Our review of literature identified both supervised [58, 59] and un-supervised [60] approaches that have been proposed for WSD.

Our AR analysis also indicates the co-occurrence of corpus analysis and semantic annotation. As stated by Ng and Zelle [61], a corpus is a resource and a technique that greatly benefits NLP tasks related to text mining and semantic annotation. Upon closer investigation of literature, we identify many research that utilized corpus for automatic discovery of concepts and relations among them using semantic annotation. Freitas et al. [62] employed semantic annotation with Wikipedia to investigate a natural language query mechanism for Linked Data. Comeau [63] used the National Center for Biotechnology Information disease corpus to annotate BioCreative collection.

In the recent years, ontology based inference and reasoning have also received much attention. Semantic web technologies, specifically ontologies, semantic annotation, and reasoning engines provide an excellent foundation for NL inference [64, 65, 66]. Our AR analysis confirms this notion, and suggests a strong association between NL inference and semantic annotation in published NLP research. Ontologies are considered the backbone of the semantic web, and provide vocabulary standardization to facilitate knowledge acquisition and processing [66]. The concepts and relationships represented as ontologies may be parsed via NLP techniques. For example, Ruiz-Martinez et al. Maria Ruiz-Martinez et al. [66] used NLP processing techniques to obtain relevant concepts and relations from biomedical text to be included in an ontology. With the advancement in semantic web techniques,

using semantic annotation for NL inference is prime for further research.

Our AR mining analysis also indicates the co-occurrence of IE and IR with the text generation. As discussed in section 2, text generation is one of the two branches of activities in NLP. One of the important subtasks in text generation is determining what kind of information should be extracted or retrieved, and then communicating in the text [8], which involves content determination and discourse structuring, both are close related to IE and IR tasks. For example, Liu et al. [29] designed and implemented a computer assisted writing system, where IE and IR techniques were used to retrieve keywords from the document corpus, and then used as input for the text generation. Similarly, Liu et al. [67] designed an intelligent computer assisted blog writing system utilizing IR techniques to obtain example texts from the Web as the input for blog text generation.

6. Conclusion

In recent years, NLP technologies have gained increasing level of sophistication. Our review of IS literature indicates that NLP research is gaining maturity and credence among IS researchers, as evidenced by the increasing number of publications, as well as the breath of algorithms, tasks, and application areas covered. As an applied discipline, IS research brings its strength in building artifacts to the NLP research landscape. On the other hand, IS research also involves the investigation of social aspects of technology (as we noted before, research on the impact of NLP on human agency is still quite sparse). With maturity of NLP research on the technology side, there is an increasing need for more behavioral research (either quantitative or qualitative) to develop and justify theories that explain or predict the interplay between NLP artifacts and their social environment. Our review of literature also highlights the need for more research in speech recognition, which has the least number of publications. The advancement in speech recognition models, especially deep learning algorithms, has resulted in wide popularity of speech recognition applications. This represents prime opportunities for IS researchers to design and develop more artifacts to utilize related technologies.

Our search procedure of the NLP literature yielded multiple articles that did not contain term “natural language processing” in their title, abstract, and/or keywords although the articles were actual research on NLP [68, 69]. NLP is a clearly defined AI technique and authors of articles that conduct

research in this area may want to include “Natural Language Processing” or related terms in their title and/or keywords. Similarly, we observed studies that were listed as NLP, but does not address a related research [70]. As the field of NLP becomes more significant in IS, it is imperative that the terms we use to identify our articles are appropriate and accurate. This will ensure that studies are easily identifiable and accessible to academic and practitioner communities.

Another noteworthy evidence from our initial assessment is that existing framework reviewing IS literature (e.g., Palvia et al. [20]) do not have explicit considerations for classifying design science research. Future research will investigate how to integrate design research classification schema into the broader framework for reviewing IS literature.

7. References

- [1] E. Kumar, *Natural language processing*, IK International Pvt Ltd, 2011.
- [2] S. Pichai: Google CEO
<https://www.technologyreview.com/s/601530/google-thinks-youre-ready-to-converse-with-computers/>
- [3] J. Hirschberg and C. D. Manning, "Advances in natural language processing", *Science*, 349, 2015, pp. 261-266.
- [4] A. Leuski and D. Traum, "NPCEditor: Creating Virtual Human Dialogue Using Information Retrieval Techniques", *Ai Magazine*, 32, 2011, pp. 42-56.
- [5] V. F. Lopez, L. Alonso and M. N. Moreno, "A SOMAgent for machine translation", *Expert Systems with Applications*, 37, 2010, pp. 7993-7996.
- [6] Q. He, C. A. W. Glas, M. Kosinski, D. J. Stillwell and B. P. Veldkamp, "Predicting self-monitoring skills using textual posts on Facebook", *Computers in Human Behavior*, 33, 2014, pp. 69-78.
- [7] T. K. Fan and C. H. Chang, "Sentiment-oriented contextual advertising", *Knowledge and Information Systems*, 23, 2010, pp. 321-344.
- [8] E. Reiter, R. Dale and Z. Feng, *Building natural language generation systems*, MIT Press, 2000.
- [9] T. Winograd, "Understanding natural language", *Cognitive psychology*, 3, 1972, pp. 1-191.
- [10] J. Cuzzola, J. Jovanovic, E. Bagheri and D. Gasevic, "Evolutionary fine-tuning of automated semantic annotation systems", *Expert Systems with Applications*, 42, 2015, pp. 6864-6877.
- [11] A. M. Azmi and S. M. Alzanin, "Aara'-a system for mining the polarity of Saudi public opinion through e-newspaper comments", *Journal of Information Science*, 40, 2014, pp. 398-410.
- [12] T. Joachims, T. Hofmann, Y. Yue and C.-N. Yu, "Predicting structured objects with support vector machines", *Communications of the ACM*, 52, 2009, pp. 97-104.
- [13] M. Ptaszynski and Y. Momouchi, "Part-of-speech tagger for Ainu language based on higher order Hidden

- Markov Model", *Expert Systems with Applications*, 39, 2012, pp. 11576-11582.
- [14] L.-C. Yu, J.-L. Wu, P.-C. Chang and H.-S. Chu, "Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news", *Knowledge-Based Systems*, 41, 2013, pp. 89-97.
- [15] H. Mangassarian and H. Artail, "A general framework for subjective information extraction from unstructured English text", *Data & Knowledge Engineering*, 62, 2007, pp. 352-367.
- [16] G. H. Grant and S. J. Conlon, "EDGAR extraction system: an automated approach to analyze employee stock option disclosures", *Journal of Information Systems*, 20, 2006, pp. 119-142.
- [17] R. Lopez-Cozar, Z. Callejas and D. Griol, "Using knowledge of misunderstandings to increase the robustness of spoken dialogue systems", *Knowledge-Based Systems*, 23, 2010, pp. 471-485.
- [18] R. Valencia-Garcia, R. Martinez-Bejar and A. Gasparetto, "An intelligent framework for simulating robot-assisted surgical operations", *Expert Systems with Applications*, 28, 2005, pp. 425-433.
- [19] J. Webster and R. T. Watson, "Analyzing the past to prepare for the future: Writing a", *MIS quarterly*, 26, 2002, pp. 13-23.
- [20] P. Palvia, E. Mao, A. Salam and K. S. Soliman, "Management information systems research: what's there in a methodology?", *Communications of the Association for Information Systems*, 11, 2003, pp. 16.
- [21] H. Leopold, R.-H. Eid-Sabbagh, J. Mendling, L. G. Azevedo and F. A. Baiao, "Detection of naming convention violations in process models for different languages", *Decision Support Systems*, 56, 2013, pp. 310-325.
- [22] T. R. Gruber, "A translation approach to portable ontology specifications", *Knowledge acquisition*, 5, 1993, pp. 199-220.
- [23] Y. Wilks, "A preferential, pattern-seeking, semantics for natural language inference", *Artificial intelligence*, 6, 1975, pp. 53-74.
- [24] A. Kiryakov, B. Popov, I. Terziev, D. Manov and D. Ognyanoff, "Semantic annotation, indexing, and retrieval", *Web Semantics: Science, Services and Agents on the World Wide Web*, 2, 2004, pp. 49-79.
- [25] A. R. Hevner, S. T. March, J. Park and S. Ram, "Design science in information systems research", *MIS quarterly*, 28, 2004, pp. 75-105.
- [26] A. R. Hevner, "A three cycle view of design science research", *Scandinavian journal of information systems*, 19, 2007, pp. 4.
- [27] K. Demirtas, N. K. Cicekli and I. Cicekli, "Automatic categorization and summarization of documentaries", *Journal of Information Science*, 36, 2010, pp. 671-689.
- [28] O. Arazy and C. Woo, "Enhancing information retrieval through statistical natural language processing: A study of collocation indexing", *Mis Quarterly*, 31, 2007, pp. 525-546.
- [29] C.-L. Liu, C.-H. Lee, S.-H. Yu and C.-W. Chen, "Computer assisted writing system", *Expert Systems with Applications*, 38, 2011, pp. 804-811.
- [30] C.-E. Wu and R. T.-H. Tsai, "Using relation selection to improve value propagation in a ConceptNet-based sentiment dictionary", *Knowledge-Based Systems*, 69, 2014, pp. 100-107.
- [31] R. Lopez-Cozar, "Using knowledge on word-islands to improve the performance of spoken dialogue systems", *Knowledge-Based Systems*, 88, 2015, pp. 223-243.
- [32] S. Bird, E. Klein and E. Loper, *Natural language processing with Python*, "O'Reilly Media, Inc.", 2009.
- [33] M. T. Mills and N. G. Bourbakis, "Graph-based methods for natural language processing and understanding—a survey and analysis", *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, 44, 2014, pp. 59-71.
- [34] A. Montoyo, P. Martinez-Barco and A. Balahur, "Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments", *Decision Support Systems*, 53, 2012, pp. 675-679.
- [35] S. Karimi, F. Scholer and A. Turpin, "Machine Transliteration Survey", *Acm Computing Surveys*, 43, 2011.
- [36] W. Wong, W. Liu and M. Bennamoun, "Ontology Learning from Text: A Look Back and into the Future", *Acm Computing Surveys*, 44, 2012.
- [37] S. Purao, "Design research in the technology of information systems: Truth or dare", *GSU Department of CIS Working Paper*, 2002, pp. 45-77.
- [38] D. Vilares, M. A. Alonso and C. Gómez-Rodríguez, "A linguistic approach for determining the topics of Spanish Twitter messages", *Journal of Information Science*, 2014, pp. 0165551514561652.
- [39] M. Yazdani and A. Popescu-Belis, "Computing text semantic relatedness using the contents and links of a hypertext encyclopedia", *Artificial Intelligence*, 194, 2013, pp. 176-202.
- [40] G. Yang, D. Wen, Kinshuk, N.-S. Chen and E. Sutinen, "A novel contextual topic model for multi-document summarization", *Expert Systems with Applications*, 42, 2015, pp. 1340-1352.
- [41] A. Wierzbicki, T. Kaszuba, R. Nielek, P. Adamska and A. Datta, "Improving computational trust representation based on Internet auction traces", *Decision Support Systems*, 54, 2013, pp. 929-940.
- [42] Y. Zhang, J. Wen, X. Wang and Z. Jiang, "Semi-supervised learning combining co-training with active learning", *Expert Systems with Applications*, 41, 2014, pp. 2372-2378.
- [43] E. Al - Shawakfa, A. Al - Badarneh, S. Shatnawi, K. Al - Rabab'ah and B. Bani - Ismail, "A comparison study of some Arabic root finding algorithms", *Journal of the American society for information science and technology*, 61, 2010, pp. 1015-1024.
- [44] V. Rus, S. D'Mello, X. Hu and A. C. Graesser, "Recent Advances in Conversational Intelligent Tutoring Systems", *Ai Magazine*, 34, 2013, pp. 42-54.
- [45] S. Soderland, "Learning information extraction rules for semi-structured and free text", *Machine learning*, 34, 1999, pp. 233-272.
- [46] C. D. Manning, P. Raghavan and H. Schütze, *Introduction to information retrieval*, Cambridge university press Cambridge, 2008.

- [47] J. Atkinson and V. Bull, "A multi-strategy approach to biological named entity recognition", *Expert Systems with Applications*, 39, 2012, pp. 12968-12974.
- [48] M. G. Armentano, D. Godoy, M. Campo and A. Amandi, "NLP-based faceted search: Experience in the development of a science and technology search engine", *Expert Systems with Applications*, 41, 2014, pp. 2886-2896.
- [49] M. Shamsfard and A. A. Barforoush, "Learning ontologies from natural language texts", *International journal of human-computer studies*, 60, 2004, pp. 17-63.
- [50] Y. Biletskiy, J. A. Brown and G. Ranganathan, "Information extraction from syllabi for academic e-Advising", *Expert Systems with Applications*, 36, 2009, pp. 4508-4516.
- [51] M. Andres Paredes-Valverde, M. Angel Rodriguez-Garcia, A. Ruiz-Martinez, R. Valencia-Garcia and G. Alor-Hernandez, "ONLI: An ontology-based system for querying DBpedia using natural language paradigm", *Expert Systems with Applications*, 42, 2015, pp. 5163-5176.
- [52] D. Milne and I. H. Witten, "An open-source toolkit for mining Wikipedia", *Artificial Intelligence*, 194, 2013, pp. 222-239.
- [53] V. Dufour-Lussier, F. Le Ber, J. Lieber and E. Nauer, "Automatic case acquisition from texts for process-oriented case-based reasoning", *Information Systems*, 40, 2014, pp. 153-167.
- [54] J. A. Bateman, J. Hois, R. Ross and T. Tenbrink, "A linguistic ontology of space for natural language processing", *Artificial Intelligence*, 174, 2010, pp. 1027-1071.
- [55] R. E. Vlas and W. N. Robinson, "Two rule-based natural language strategies for requirements discovery and classification in open source software development projects", *Journal of Management Information Systems*, 28, 2012, pp. 11-38.
- [56] R. Agrawal, T. Imieliński and A. Swami, "Mining association rules between sets of items in large databases", *ACM SIGMOD Record*, 22, 1993, pp. 207-216.
- [57] S. Brin, R. Motwani, J. D. Ullman and S. Tsur, *Dynamic itemset counting and implication rules for market basket data*, *ACM SIGMOD Record*, ACM, 1997, pp. 255-264.
- [58] H. Yu, W. Kim, V. Hatzivassiloglou and J. Wilbur, "A large scale, corpus-based approach for automatically disambiguating biomedical abbreviations", *Acm Transactions on Information Systems*, 24, 2006, pp. 380-404.
- [59] N. Fernandez, J. Arias Fisteus, L. Sanchez and G. Lopez, "IdentityRank: Named entity disambiguation in the news domain", *Expert Systems with Applications*, 39, 2012, pp. 9207-9221.
- [60] A. Duque, J. Martinez-Romo and L. Araujo, "Choosing the best dictionary for Cross-Lingual Word Sense Disambiguation", *Knowledge-Based Systems*, 81, 2015, pp. 65-75.
- [61] H. T. Ng and J. Zelle, "Corpus-based approaches to semantic interpretation in NLP", *AI magazine*, 18, 1997, pp. 45.
- [62] A. Freitas, J. G. Oliveira, S. O'Riain, E. Curry and J. C. P. Da Silva, *Querying linked data using semantic relatedness: a vocabulary independent approach*, *Natural Language Processing and Information Systems*, Springer, 2011, pp. 40-51.
- [63] D. C. Comeau, "Natural language processing pipelines to annotate BioC collections with an application to the NCBI disease corpus (vol 2014, bau056, 2014)", *Database-the Journal of Biological Databases and Curation*, 2014.
- [64] T. Lee, "Constraint-based ontology induction from online customer reviews", *Group Decision and Negotiation*, 16, 2007, pp. 255-281.
- [65] R. Gacitua, P. Sawyer and P. Rayson, "A flexible framework to experiment with ontology learning techniques", *Knowledge-Based Systems*, 21, 2008, pp. 192-199.
- [66] J. Maria Ruiz-Martinez, R. Valencia-Garcia, J. Tomas Fernandez-Breis, F. Garcia-Sanchez and R. Martinez-Bejar, "Ontology learning from biomedical natural language documents using UMLS", *Expert Systems with Applications*, 38, 2011, pp. 12365-12378.
- [67] C.-L. Liu, C.-H. Lee and B.-Y. Ding, "Intelligent computer assisted blog writing system", *Expert Systems with Applications*, 39, 2012, pp. 4496-4504.
- [68] C. Li, A. Sun, J. Weng and Q. He, "Tweet segmentation and its application to named entity recognition", *Knowledge and Data Engineering, IEEE Transactions on*, 27, 2015, pp. 558-570.
- [69] A. Rodrigo and A. Penas, "On Evaluating the Contribution of Validation for Question Answering", *Knowledge and Data Engineering, IEEE Transactions on*, 27, 2015, pp. 1157-1161.
- [70] N. Nistor, B. Baltes, M. Dascalu, D. Mihaila, G. Smeaton and S. Trausan-Matu, "Participation in virtual academic communities of practice under the influence of technology acceptance and community factors. A learning analytics application", *Computers in Human Behavior*, 34, 2014, pp. 339-344.